Adverse News Detection and Classification System for Anti-Money Laundering Investigations

Fabiola Rivera Rodríguez
Master in Computer Science
Advisor: Alfredo Cruz, Ph.D.
Polytechnic University of Puerto Rico
Graduate Project EXPO, May 2026

Abstract *— Adverse media screening is an important component of Anti-Money Laundering (AML) compliance, requiring the analysis of large volumes of unstructured news data. This project presents the development of an automated adverse media detection system using Natural Language Processing techniques to support AML-focused risk analysis. The system collects multilingual news articles from publicly available Really Simple Syndication feeds, applies Named Entity Recognition using spaCy models for English and Spanish, and incorporates rule-based filtering and normalization to improve entity extraction consistency. A keyword-based risk-scoring methodology assigns article-level risk scores, which are later aggregated to evaluate entity-level exposure to adverse media indicators. A Streamlit-based web application was developed to facilitate interactive exploration of ranked entities and associated articles. The system achieved 90.66% precision in entity extraction, with a 95% confidence interval ranging from 85.55% to 94.09%, demonstrating the effectiveness of the proposed approach for multilingual adverse media analysis within AML contexts.*

Keywords *— Adverse Media Screening, Anti-Money Laundering, Natural Language Processing, Named Entity Recognition.*

INTRODUCTION

Anti-Money Laundering (AML) frameworks require financial institutions to identify, assess, and continuously monitor risks related to financial crime. Regulatory organizations such as the Financial Action Task Force (FATF) emphasize the importance of ongoing monitoring and the integration of external information sources into due diligence processes [1]. Adverse media screening contributes to this objective by incorporating publicly available information, including news articles, legal reports, and online publications, to identify entities potentially linked to financial crime indicators.

Traditional adverse media screening approaches often rely on manual review processes or simple keyword-based searches. Although these approaches can identify relevant information, they often struggle to efficiently process large volumes of multilingual, unstructured textual data. In addition, many existing approaches focus primarily on document-level analysis without explicitly establishing relationships between extracted entities and aggregated risk indicators across multiple articles.

The increasing volume of digital news content and the multilingual nature of adverse media sources create additional challenges for AML monitoring processes. Financial institutions and compliance teams must review large amounts of unstructured information originating from different jurisdictions, languages, and publication formats. Manual review procedures are often time-consuming and difficult to scale, while traditional keyword-based methods may generate excessive false positives or fail to capture contextual relationships between entities and risk indicators. These limitations highlight the need for interpretable, scalable systems that can organize adverse media information into structured, actionable outputs.

Natural Language Processing (NLP) techniques provide opportunities to improve adverse media analysis by automating the extraction and organization of relevant information from textual sources. Named Entity Recognition (NER) enables the identification of individuals, organizations, and locations associated with adverse media indicators. However, challenges remain in integrating multilingual processing, explainable risk scoring, and entity-level aggregation into a unified framework aligned with AML requirements [2] and NLP-related analytical challenges identified in prior research [3, 4].

This project presents the development of an automated adverse media detection system designed to support AML-focused risk analysis through multilingual news processing and entity-level risk aggregation. The system combines RSS-based news collection, multilingual NER, rule-based filtering, article-level risk scoring, and entity-level aggregation within a unified and interpretable analytical framework. A Streamlit-based web application was also developed to facilitate interactive exploration of ranked entities and associated article-level risk information [5].

Literature Review

Recent research in financial crime detection has highlighted the growing use of Artificial Intelligence and NLP techniques to support AML processes and automate the analysis of unstructured textual data. NLP-based methods have demonstrated potential to improve the detection of financial crime indicators across large volumes of news articles, legal reports, and publicly available online content [3, 4].

NER is one of the most widely used NLP techniques in adverse media analysis. NER models identify and classify entities such as individuals, organizations, and locations from textual sources. In multilingual environments, NER systems face additional challenges related to language structure, contextual ambiguity, and variations in naming conventions. These factors may affect extraction consistency and reduce the reliability of downstream analytical processes [6, 7]. Many adverse media screening systems rely on keyword matching, document classification, or single entity extraction techniques. These methods can identify potentially relevant information, yet they may struggle to capture contextual relationships between entities and financial crime indicators across multiple articles. In addition, several machine learning-based frameworks operate as “black-box” systems, limiting interpretability and reducing transparency within regulated AML environments where explainable analytical outputs are essential [3, 4].

Previous studies have also highlighted limitations associated with multilingual adverse media analysis and entity-level aggregation. Many existing systems focus primarily on document-level classification without incorporating mechanisms to aggregate article-level risk information into interpretable entity-level indicators. In addition, localized geographic references and multilingual sources may introduce inconsistencies that affect extraction accuracy and downstream risk evaluation processes [8, 6].

These limitations highlight the need for an interpretable, scalable framework that integrates multilingual entity extraction, explainable risk scoring, and entity-level aggregation into a unified adverse media screening process aligned with AML requirements. The proposed system addresses these challenges through a modular pipeline that combines multilingual NLP techniques, rule-based filtering, weighted keyword scoring, and entity-level risk aggregation within a unified analytical framework.

Methodology

The proposed system follows a modular pipeline composed of five primary components: data collection, data preprocessing, entity extraction, article-level risk scoring, and entity-level aggregation. Together, these components transform multilingual news data into structured analytical outputs that facilitate adverse media evaluation, entity prioritization, and AML-focused analysis.

News articles are collected from publicly available RSS feeds that include local, international, financial, and official government sources. The collected data includes article titles, summaries, publication dates, source names, and URLs. The dataset is stored in CSV and Parquet formats to facilitate processing, validation, and reproducibility. Data preprocessing standardizes the collected information by removing duplicate records, correcting encoding issues, normalizing text fields, and generating unique article identifiers through hashing. These preprocessing operations improve data consistency and reduce variations that could negatively affect downstream entity extraction and risk evaluation tasks.

The overall architecture of the proposed adverse media detection pipeline is illustrated in Figure 1.

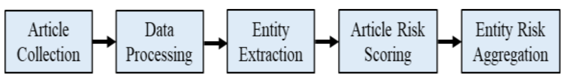

Figure 1
Overall Pipeline of the Proposed Adverse Media Screening System

ENTITY EXTRACTION

NER is performed using spaCy pre-trained models for both English and Spanish text [9]. The system extracts entities associated with individuals, organizations, and locations using the PERSON, ORG, GPE, and LOC labels. Rule-based filtering and normalization operations are applied to remove noise, validate entities, and standardize naming variations.

The system design allows entity extraction across both local and international news sources. Additional rule-based enhancements improve extraction consistency and reduce classification ambiguity within the dataset. These enhancements include filtering invalid entities, resolving classification conflicts, validating location references, and excluding government or regulatory organizations that primarily function as reporting entities rather than subjects of risk analysis.

Normalization and standardization tasks are applied to reduce duplication caused by formatting differences and naming variations across sources. Location entities are further validated using predefined country and Puerto Rico-specific geographic reference lists. Duplicate entities within the same article are removed to prevent frequency inflation during downstream aggregation.

Risk Scoring

The risk-scoring component evaluates article content to identify potential financial crime indicators associated with adverse media reporting. Risk evaluation is performed using normalized text derived from article titles and summaries, allowing the system to analyze both headline-level and contextual information.

The proposed methodology applies predefined financial crime-related keyword categories composed of strong and moderate indicators. Strong indicators represent terms more directly associated with financial crime activity, whereas moderate indicators capture broader contextual references. A weighted scoring technique assigns higher values to strong keyword matches and lower values to moderate matches, generating an aggregated article-level risk score.

Articles are subsequently classified into Low Risk, Medium Risk, and High Risk categories based on the calculated score. This scoring methodology was designed to maintain interpretability by allowing risk classifications to be directly associated with identifiable keyword indicators. The resulting article-level scores serve as the foundation for downstream entity-level aggregation and prioritization.

Entity-Level Aggregation

After article-level scoring is complete, extracted entities are linked to their associated article-level scores via a unique article identifier. Each entity inherits the scores of the articles in which it appears, allowing risk indicators to accumulate across multiple occurrences.

Aggregation steps compute entity-level metrics, including article frequency, total risk score, and average risk score. These metrics are normalized and combined to produce an entity risk index that ranks entities by their exposure to adverse media indicators.

This aggregation process transforms single article evaluations into broader entity-level analytical indicators that support prioritization and investigative review. By concentrating risk indicators across multiple articles, the proposed approach facilitates the identification of individuals and organizations that appear repeatedly in higher-risk adverse media content.

Web Application

A Streamlit-based web application [5] was developed to facilitate interactive exploration of the system outputs. The dashboard allows users to filter entities by type, search for specific entities, and review article-level scoring information associated with ranked entities. The interface also incorporates filtering and ranking capabilities that support focused exploration of adverse media indicators across multilingual news sources.

The web application demonstrates the practical applicability of the proposed framework by transforming analytical outputs into an accessible and interactive interface. Figure 2 presents the main dashboard view, including entity ranking and filtering to explore aggregated risk results.

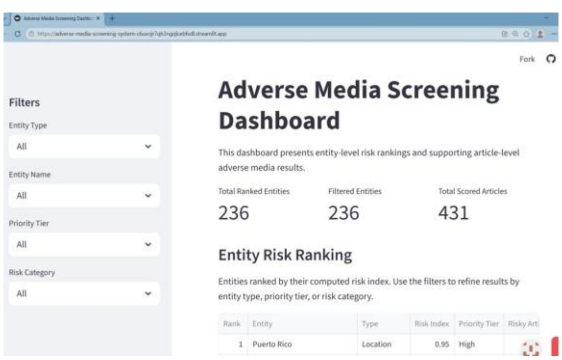

Figure 2
Adverse Media Screening Dashboard with Entity Ranking and Sidebar Filters

RESULTS

In this section, the results of the entity extraction process, risk scoring, entity-level analysis, and system performance are presented.

Entity Extraction Performance

The entity extraction process was evaluated using precision as the primary metric. A manually reviewed random sample of 182 extracted entities was used for validation. Based on this evaluation, the system achieved a precision of 90.66%, with a 95% confidence interval ranging from 85.55% to 94.09% calculated using the Wilson score method [10].

The findings suggest that most extracted entities correctly represent valid individuals, organizations, or locations within the analyzed articles. The relatively narrow confidence interval indicates stable and dependable extraction performance across multilingual news sources. Maintaining high precision is particularly important in AML-focused adverse media analysis, where excessive false positives can undermine the effectiveness and interpretability of downstream risk evaluation tasks.

Risk Scoring Results

The risk-scoring component was applied to the processed dataset to identify and quantify financial crime-related indicators within news articles. The weighted scoring methodology assigns higher values to strong keyword matches and lower values to moderate matches, resulting in an aggregated article-level risk score and a corresponding risk classification. This scoring approach allows the system to differentiate between direct financial crime references and broader contextual indicators while maintaining transparency through explicit keyword associations.

Most articles were classified as Low Risk or Medium Risk, while a smaller subset was classified as High Risk. This distribution reflects the expected behavior of adverse media content, where explicit financial crime indicators appear less frequently than broader contextual references. Figure 3 presents the distribution of article classifications across risk levels. The observed distribution also reflects the broader nature of adverse media reporting, where contextual references and indirect associations occur more frequently than explicit mentions of financial crime activity.

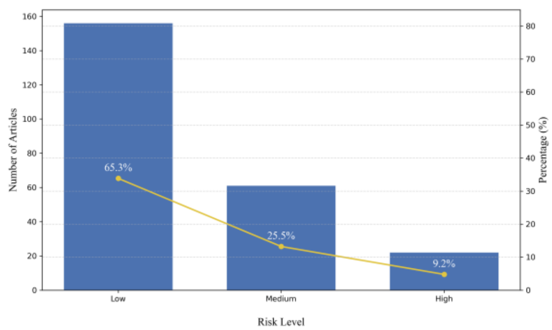

Figure 3
Distribution of Articles by Risk Level

The analysis shows that moderate indicators occurred more frequently across the dataset, while strong indicators contributed more significantly to higher-risk classifications due to their higher assigned weights. This behavior demonstrates that the scoring methodology effectively differentiates between varying levels of adverse media relevance, maintaining interpretability through explicit keyword-based associations.

The resulting article-level scores provide the analytical foundation for downstream entity-level aggregation and prioritization.

Entity-Level Results

Following the risk-scoring process, article-level scores were associated with extracted entities to enable entity-level analysis. Each entity inherits the scores of the articles in which it appears, allowing risk exposure to accumulate across multiple occurrences. Entities that appear more frequently in elevated-risk articles accumulate greater overall exposure, thereby becoming more prominent in the ranking process. This aggregation behavior enables the methodology to identify recurring patterns of adverse media exposure that may not be visible in individual article evaluations.

The aggregated entity dataset provides a ranked view of individuals and organizations according to their calculated entity risk index values. Figure 4 presents the highest-ranked entities identified within the analyzed dataset. The ranking process facilitates focused analysis by highlighting entities that repeatedly appear in articles with stronger adverse media indicators and higher aggregated risk.

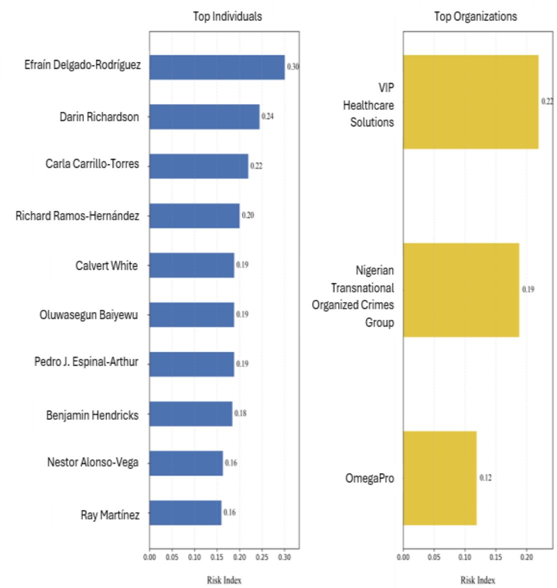

Figure 4
Highest-Ranked Individuals and Organizations by Aggregated Risk Index

A relatively small subset of entities concentrates higher risk scores, reflecting repeated associations with adverse media indicators across multiple articles. Individuals occupied most of the highest-ranking positions, while organizations appeared less frequently among the top-ranked entities. This behavior reflects the nature of financial crime reporting, where individuals are more commonly identified as primary subjects of investigations and enforcement actions.

These findings demonstrate how entity-level aggregation transforms isolated article evaluations into broader analytical indicators that support prioritization and investigative review in AML-focused adverse media analysis.

System Performance

The entity ranking results highlight the individuals and organizations associated with higher risk; it is also important to examine how scores are distributed across the broader set of extracted entities. This analysis provides additional insight into the behavior and interpretability of the proposed scoring methodology.

Entity risk scores were analyzed by entity type, including individuals, organizations, and locations. Most entities exhibited relatively low scores, with only a limited number accumulating substantially higher values. Figure 5 presents the distribution of entity risk scores across entity categories.

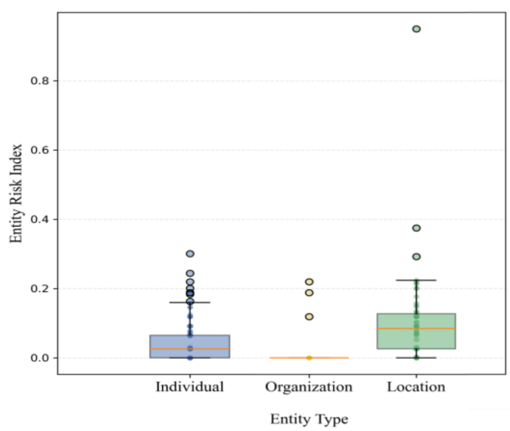

Figure 5
Distribution of Entity Risk Index Values by Entity Type

Certain location entities exhibited elevated scores due to repeated associations with higher-risk articles rather than direct attribution of financial crime involvement. This behavior highlights the contextual nature of location-based adverse media references, in which geographic entities may appear frequently within risk-related reporting without necessarily representing investigative subjects themselves.

The concentration of higher scores among a relatively small subset of entities demonstrates that the proposed framework effectively differentiates between lower- and higher-risk adverse media associations. This behavior supports the interpretability and prioritization capabilities of the entity-level aggregation approach within AML-focused analytical contexts.

Discussion

The proposed system demonstrates how multilingual NLP techniques and entity-level aggregation can be integrated into an interpretable adverse media analysis workflow aligned with AML requirements. By combining multilingual NER, rule-based refinement, weighted keyword scoring, and entity-level aggregation, the methodology transforms unstructured news content into structured analytical outputs that support risk-focused evaluation and investigative review.

The entity-centric design provides practical value within adverse media monitoring environments by concentrating risk exposure across multiple articles and highlighting entities repeatedly associated with elevated-risk reporting. In contrast to standalone document-level analysis, the aggregation methodology facilitates broader analytical interpretation and supports review processes commonly associated with EDD activities. The use of interpretable, keyword-based scoring also improves transparency by directly linking article classifications to identifiable indicators.

The modular structure of the system additionally facilitates future adaptation and integration into evolving AML monitoring environments, where scalability, analytical transparency, and multilingual processing remain important operational considerations.

Despite these contributions, several limitations remain. The methodology relies on predefined keyword categories and rule-based refinement operations, which may not capture all contextual variations present in multilingual news content. In addition, the use of article titles and summaries instead of full article text may limit contextual depth during entity extraction and risk evaluation. Variations in multilingual reporting styles, naming conventions, and localized references may also affect extraction consistency across sources.

Future Work

Future work may focus on extending the proposed system as a filtering and prioritization component that supports broader adverse media monitoring workflows. One important direction is to integrate ongoing data storage and incremental article processing to support automated updates while avoiding reprocessing previously analyzed content.

Additional improvements may include entity validation, classification feedback, and case review tracking within the web application. Future enhancements may also strengthen entity resolution to reduce fragmentation caused by multilingual references and naming variations across sources.

Evaluation can be further expanded through additional performance metrics, expert validation, and comparisons with baseline keyword-based approaches to provide a more comprehensive assessment of system effectiveness.

Conclusion

This project presented the development of an automated adverse media detection system designed to support AML risk analysis through multilingual news processing and entity-level aggregation. The proposed methodology integrates multilingual NER, rule-based refinement, weighted keyword-based risk scoring, and entity-level aggregation into a unified and interpretable analytical workflow.

The findings demonstrate that the proposed approach can effectively transform unstructured news content into structured risk indicators that support adverse media evaluation and entity prioritization. The Streamlit-based web application further demonstrates the operational value of the system through interactive exploration of ranked entities and associated article-level risk information.

Overall, the proposed system provides an interpretable and scalable approach for multilingual adverse media analysis aligned with AML-focused monitoring and EDD activities.

References

[1] Financial Action Task Force. "FATF Recommendations." FATF-GAFI, 2024. [Online]. Available: https://www.fatf-gafi.org/en/topics/fatf-recommendations.html.

[2] Financial Crimes Enforcement Network (FinCEN). "Anti-Money Laudering Act of 2020." U.S. Department of the Treasury, 2020. [Online]. Available: https://www.fincen.gov/resources/statutes-and-regulations/anti-money-laundering-act-2020.

[3] S. Roy. "AI Adoption to Combat Financial Crime: Study on Natural Language Processing in Adverse Media Screening of Financial Services in English and Bangla multilingual interpretation." *arXiv (Cornell University)*, 2024, doi:10.48550/arxiv.2412.12171.

[4] D. Khurana, A. Koli, K. Khatter, and S. Singh. "Natural Language processing: State of the art, Current Trends and Challenges." *Multimedia Tools and Applications*, vol. 82, no. 3, 3713–3744, 2022, doi: 10.1007/s11042-022-13428-4.

[5] Streamlit, Inc., "Streamlit: A faster way to build and share data apps." Streamlit.io, 2025. [Online]. Available: https://streamlit.io/

[6] T. Chavan and S. Patil. "Named Entity Recognition (NER) for news articles." *International Journal of Advanced Research in Engineering and Technology*, vol. 2, no. 1, 103-112, 2024, doi:10.34218/ijaird.2.1.2024.10.

[7] Warto et al. "Systematic Literature Review on Named Entity Recognition: Approach, Method, and Application." *Statistics, Optimization & Information Computing*, vol. 12, no. 4, 907–942, 2024, doi: 10.19139/soic-2310-5070-1631.

[8] K. Boudt, O. Delmarcelle, and P. Ringoot. "A news monitoring system to detect relevant news for the anti-money laundering supervision of financial institutions." *Risk Sciences*, vol. 1, 100018, 2025. Doi: 10.1016/j.risk.2025.100018.

[9] spaCy. "Industrial-Strength Natural Language Processing in Python." spaCy, 2015. Available: https://spacy.io/

[10] E. B. Wilson. "Probable Inference, the Law of Succession, and Statistical Inference." *Journal of the American Statistical Association*, vol. 22, no. 158, 209–212, 1927. Doi: 10.1080/01621459.1927.10502953.